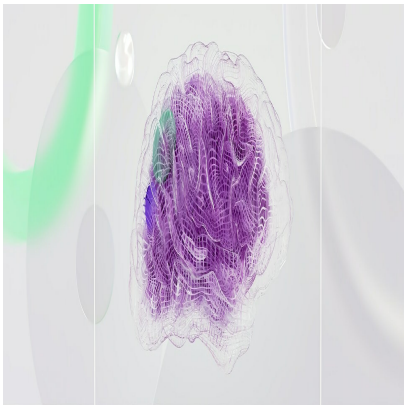

Intelligenza artificiale: un codice etico per lo sviluppo umano

Descrizione

Il 2023 Ã¨ stato lâ??anno in cui lâ??Intelligenza Artificiale (IA) ha segnato una significativa e collettiva diffusione nei piÃ¹ svariati ambiti del sociale. Si delineano nuove prospettive e pressanti interrogativi. Basta considerare che la progressione degli investimenti globali sullâ??IA prevede per il 2024 investimenti per circa 300 milioni di dollari e oltre 420 per il 2025. Per il 2030, ben 1900 miliardi di dollari. E poi lâ??utilizzo dellâ??IA in ambito bellico e la manipolazione del consenso nelle prossime elezioni in Europa e negli Stati Uniti. Si delineano, cosÃ¬, nuove strutture di potere politico.

Intelligenza artificiale, fraternitÃ e pace Ã¨ stato il tema chiave di Papa Francesco nel Messaggio per la Giornata mondiale della pace 2024. Ã? stata tracciata la strada da seguire per un codice etico. Â«Lâ??intelligenza artificiale diventerÃ sempre piÃ¹ importante. Le sfide che pone sono tecniche, ma anche antropologiche, educative, sociali e politiche. Occorre essere consapevoli delle rapide trasformazioni in atto e gestirle in modo da salvaguardare i diritti umani fondamentali, rispettando le istituzioni e le leggi che promuovono lo sviluppo umano integrale. Lâ??intelligenza artificiale dovrebbe essere al servizio del migliore potenziale umano e delle nostre piÃ¹ alte aspirazioni, non in competizione con essi.Â»

Lâ??esortazione di Papa Francesco Ã¨ volta a una responsabilizzazione comunitaria. Un percorso che dovrebbe avere tre obiettivi sostanziali: consapevolezza nellâ??abitare le innovazioni tecnologiche; interdisciplinaritÃ delle competenze; regolamentazioni, secondo principi condivisi di etica sociale, dellâ??innovazione tecnologica nellâ??ottica dello sviluppo umano-centrico.

Obiettivi, questi, certo non facili da raggiungere tuttavia ineludibili. Da un lato il riduzionismo antropologico del paradigma tecnocratico, dominio degli algoritmi o algocrazia. Dallâ??altro lo sviluppo umano-centrico (*human-centric /human-centred*) che tiene conto delle innovazioni tecnologiche ma che sa coniugarle secondo unâ??etica degli algoritmi ovvero Â«algoreticaÂ», secondo il puntuale neologismo coniato da P. Paolo Benanti. Per un umanesimo digitale che Â«non trasforma lâ??essere umano in una macchina e non interpreta le macchine come esseri umani. Che riconosce la peculiaritÃ dellâ??essere umano e delle sue capacitÃ , servendosi delle tecnologie digitali per ampliarle, non per restringerle.Â»

Molteplici i principi etici proposti. Come richiamato da Luciano Floridi, Â«lâ??enorme volume di principi proposti â?? piÃ¹ di 160 nel 2020, secondo lâ??*AI Ethics Guidelines Global Inventory* (la rassegna globale delle linee guida etiche per lâ??IA) di *Algorithm Watch* â?? rischia di diventare soverchiante e fuorviante, sollevando due potenziali problemi. O i vari insiemi di principi etici per lâ??IA sono simili, portando a inutili ripetizioni e ridondanze, oppure, se differiscono in modo significativo, sono suscettibili di ingenerare confusione e ambiguitÃ . Il peggior risultato sarebbe quello di creare un â??mercato di principiâ?• in cui le parti interessate potrebbero essere tentate di â??acquistareâ?• quelli piÃ¹ allettanti.Â»

Allora come risolvere il problema della â??proliferazione dei principiâ?•? Sempre Floridi, sulla base di unâ??analisi comparativa, identifica Â«un quadro generale costituito da cinque principi fondamentali per lâ??IA etica. Prendendo atto dei limiti e valutando le implicazioni di questo quadro etico per gli sforzi futuri volti a creare leggi, regole, standard e le migliori pratiche (*best practices*) per lâ??IA etica in unâ??ampia varietÃ di contesti.Â» I cinque principi individuati sono: beneficenza, non maleficenza, giustizia, autonomia ed esplicabilitÃ . I primi quattro ordinano, classificano e raggruppano norme etiche che, come da consolidate riflessioni bioetiche, necessitano di contenuto, specificitÃ e gerarchizzazione. Il quinto e nuovo principio di esplicabilitÃ Ã¨ da intendere Â«include sia il senso epistemologico di intelligibilitÃ *â??* come risposta alla domanda: come funziona? â?? sia quello etico di responsabilitÃ (*accountability*) come risposta alla domanda: chi Ã¨ responsabile del modo in cui funziona?Â»

Varie istituzioni, nazionali e internazionali, stanno sviluppando percorsi con la finalitÃ di regolamentazioni condivise, pur nella consapevolezza della complessitÃ del tema e dei rilevantissimi interessi coinvolti. I procedimenti di regolamentazione in corso significano processi di democratizzazione. Evitando, cosÃ¬, che la c.d. â??scatola neraâ?• (*black box*), ovvero lâ??inesplicabilitÃ di modelli di IA â?? rispetto a modelli di IA interpretabili e spiegabili â?? possa rappresentare un alibi per lâ??accettazione passiva di qualsiasi procedura o risultato. Accantonando, erroneamente, lâ??esplicabilitÃ ovvero i chiarimenti sulle cause delle decisioni (*output*), i meccanismi di elaborazione e interpretazione dei dati, la valutazione dellâ??appropriatezza o inadeguatezza dei risultati prodotti. La strada da seguire significa coniugare dimensione umana e dimensione artificiale, favorendo lâ??incontro tra la fiducia e la cautela.

Ecco la necessitÃ di governare con regolamentazioni, abitando le innovazioni tecnologiche nella consapevolezza e cooperando nellâ??interdisciplinaritÃ delle competenze.

Citiamo, in particolare, i piÃ¹ recenti documenti pubblicati alla fine del 2023. In particolare, il Decalogo del Garante per la Protezione dei Dati Personali (GPDP); lâ??*Artificial Intelligence Act* dellâ??Unione europea (EU AI Act); il Rapporto intermedio (*Governing AI for Humanity*) dellâ??organo consultivo del Segretario generale delle Nazioni Unite su rischi, opportunitÃ e governance internazionale dellâ??IA.

Il GPDP ha emanato nel settembre 2023, il Decalogo per la realizzazione di servizi sanitari nazionali attraverso sistemi di IA. Sono richiamati tre principi: conoscibilitÃ , non esclusivitÃ della decisione algoritmica e non discriminazione algoritmica. Per quanto riguarda la conoscibilitÃ , lâ??interessato ha il diritto di conoscere lâ??esistenza di processi decisionali basati su trattamenti automatizzati e, in tal caso, di ricevere informazioni significative sulla logica utilizzata, sÃ¬ da poterla comprendere. Per quanto riguarda la non esclusivitÃ della decisione algoritmica, nel processo decisionale deve esistere un intervento umano capace di controllare, validare ovvero smentire la decisione automatica (c.d.

human in the loop). Infine, la non discriminazione algoritmica. Il titolare del trattamento utilizzi sistemi di IA affidabili che riducano le opacitÃ , gli errori dovuti a cause tecnologiche e/o umane. Verificando periodicamente lâ??efficacia anche alla luce della rapida evoluzione delle tecnologie impiegate, delle procedure matematiche o statistiche appropriate per la profilazione, mettendo in atto misure tecniche e organizzative adeguate.

Lâ??Unione europea con lâ??Artificial Intelligence Act (EU AI Act) si Ã¨ prefisso come obiettivo assicurare che i cittadini europei possano beneficiare delle nuove tecnologie in conformitÃ ai valori, ai diritti fondamentali e ai principi dellâ??Unione. Non Ã¨ stato ancora concluso il suo iter formale. Il Consiglio europeo ha adottato lâ??orientamento generale sulla Proposta della Commissione. Successivamente Commissione europea, Consiglio dellâ??Unione europea, Parlamento (Trilogo) hanno approvato lâ??AI Act nel dicembre 2023. Si prevede che la normativa possa essere applicativa nel 2026, ossia 24 mesi dalla formale entrata in vigore del regolamento. Un termine critico perchÃ© troppo spostato in avanti, visto il rapidissimo evolvere delle innovazioni tecnologiche.

In sintesi, lâ??accordo vieta sistemi di IA che presentano un livello di rischio inaccettabile per la sicurezza delle persone, come quelli utilizzati per il punteggio sociale (classificare le persone in base al loro comportamento sociale o alle loro caratteristiche personali). Divieti, poi, sugli usi intrusivi e discriminatori dellâ??IA. Tra questi, lâ??uso di sistemi di identificazione biometrica remota â??in tempo realeâ?• in spazi accessibili al pubblico; sistemi di categorizzazione biometrica basati su caratteristiche sensibili (ad esempio genere, razza, etnia, cittadinanza, religione, orientamento politico); sistemi di polizia predittiva (basati su profilazione, ubicazione o comportamenti criminali passati); sistemi di riconoscimento delle emozioni; estrazione non mirata di dati biometrici da Internet o da filmati di telecamere a circuito chiuso per creare database di riconoscimento facciale (in violazione dei diritti umani e del diritto alla *privacy*).

Lâ??ONU, a sua volta, tramite lâ??organo consultivo sullâ??IA del Segretario generale delle Nazioni Unite, ha lanciato nel dicembre 2023 il Rapporto intermedio â??*Governing AI for Humanity*â?•. Titolo significativo ed evocativo. Elemento centrale Ã¨ la proposta di governance dellâ??IA, individuando i principi di riferimento. InclusivitÃ : tutti i cittadini dovrebbero essere in grado di accedere e utilizzare in modo significativo gli strumenti di IA. Interesse pubblico: la governance dovrebbe andare oltre il principio di non nuocere e definire un quadro di responsabilitÃ piÃ¹ ampio per le aziende che costruiscono, distribuiscono e controllano lâ??IA, nonchÃ© per gli utenti. CentralitÃ della governance dei dati: la governance dellâ??IA non puÃ² essere separata dalla governance dei dati. Universale, interconnessa e multilaterale: la governance dellâ??IA dovrebbe dare prioritÃ al consenso universale da parte dei paesi e delle parti interessate nonchÃ© dovrebbe sfruttare le istituzioni esistenti attraverso un approccio di rete. Infine, il Diritto internazionale: la governance dellâ??IA deve essere ancorata alla Carta delle Nazioni Unite, al diritto internazionale sui diritti umani e agli obiettivi di sviluppo sostenibile. Lâ??organo consultivo sullâ??IA ha promosso per i prossimi mesi consultazioni aperte con tutte le parti interessate â?? individui, gruppi e organizzazioni â?? cosÃ¬ da offrire *feedback* prima della stesura definitiva del Rapporto.

In definitiva, il comune denominatore, costituito da consapevolezza-interdisciplinaritÃ - regolamentazione, rappresenta la prospettiva per un responsabile sviluppo umano alla luce delle rapidissime innovazioni tecnologiche. NÃ© Ã¨ plausibile, come postulato dai tecnofobi secondo un improponibile neo-luddismo, opporsi alle innovazioni tecnologiche. Ricordava Benjamin Franklin: Â« *When you are finished changing, you are finished*Â».

Aggiornamento del 20 gennaio 2024 â?? Il 18 gennaio 2024 la World Health Organization ha pubblicato la Guida: â??Ethics and governance of artificial intelligence for healthâ?•.

Si identificano sia i potenziali benefici che i potenziali rischi dellâ??uso dellâ??IA nellâ??assistenza sanitaria. Sono stati definiti sei principi da prendere in considerazione nelle politiche e nelle pratiche di governi, sviluppatori e fornitori che utilizzano lâ??IA.
I principi sono:
(1) proteggere lâ??autonomia;
(2) promuovere il benessere umano, la sicurezza umana e lâ??interesse pubblico;
(3) garantire trasparenza, â??spiegabilitÃ â?• e intelligibilitÃ ;
(4) promuovere la responsabilitÃ e lâ??affidabilitÃ ;
(5) garantire inclusivitÃ ed equitÃ ;
(6) promuovere unâ??Intelligenza Artificiale sostenibile.

Lâ??OMS pubblica questa Guida per assistere gli Stati membri nella mappatura dei benefici e delle sfide associati allâ??uso dellâ??Intelligenza Artificiale generativa che si basa sui Large Language Model (LMM).
La Guida include raccomandazioni per la governance, allâ??interno delle aziende, da parte dei governi e attraverso la collaborazione internazionale, in linea con i principi guida.

Per il download della Guida, ecco il link:
https://iris.who.int/bitstream/handle/10665/375579/9789240084759-eng.pdf?sequence=1&isAllowed=y

Crediti foto DeepMind @deepmind su Unsplash

Data di creazione
17 Gennaio 2024
Autore
lucio_romano